PAPER • **OPEN ACCESS**

A new predictive model for Plants Photosynthesis Influenced by Major Climatic Conditions

To cite this article: ZS Liu *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **291** 012016

IOP Conf. Series: Earth and Environmental Science **291** (2019) 012016 doi:10.1088/1755-1315/291/1/012016

A new predictive model for Plants Photosynthesis Influenced by Major Climatic Conditions

ZS Liu[1], WY Yang[2], XL Yu[3]

[1] College of Agriculture, Shandong Agricultural University, Taian, Shandong, China
[2] College of automation, Shenyang Aerospace University, Shenyang, Liaoning, China
[3] School of Science, Beijing University of Posts and Telecommunications, Beijing, China

Email: zzc1890@hotmail.com

Abstract: Climate change, particularly global warming, is significantly affected by atmospheric CO_2 dynamics. Plant photosynthesis is capable of fixing a large amount of airborne CO_2 and converts it into vegetation biomass and thus alleviates the greenhouse effect from atmospheric CO_2. However, how climate change and climate condition impact the dynamics of plant photosynthesis is still highly uncertain. Here we combined high frequency land surface measurements of photosynthetic CO_2 fixation data and information theory to understand the casual relationship from climate drivers on the photosynthesis rate. We found that temperature and shortwave radiation dominated photosynthesis more at forest site, while precipitation dominated photosynthesis more at grass land site. More importantly, linear regression based analysis failed to identify such relationships, which confirmed the important role of information theory in identifying nonlinear relationship within a complex system.

1. Introduction

Since preindustrial era, human activities have significantly increased atmospheric greenhouse gas (e.g., carbon dioxide) concentrations, the sources of which were mainly from fossil fuel emissions, biomass burning, land use and land cover changes [1]. Positive radiative forcings induced by airborne greenhouse gases largely warmed up the earth system and led to a substantial and observable increase in surface temperature [2]. For example, the atmospheric CO2 concentration increased from 280 ppm (ppm = parts per million) before industrial evolution to 410 ppm at present day. Consequently, global mean surface temperature increased more than half a degree Celsius and will further increase in the future [3].

Although anthropogenic activities significantly enhanced atmospheric CO2 concentration, natural ecosystems are capable of absorbing a large fraction (about a half) of human induced CO2 emissions, thus effectively mitigate the increasing trends of both greenhouse gas and surface temperature [1]. Terrestrial ecosystem (mainly forest and grass) sequestrates atmospheric CO2 molecules through photosynthesis, a biochemical reaction that combines CO2 and water and store as carbohydrate in the vegetation biomass [4]. The photosynthesis reaction is fueled by light, and significantly controlled by the activity of RuBisCo enzyme, which is highly temperature sensitive. Therefore, theatrically, terrestrial ecosystem photosynthetic capacity is co-controlled by substrates (CO2 and water), temperature (enzyme activity), and solar radiation (energy supply). However, which one is the dominant controlling factor of photosynthesis reaction is still highly uncertain for different vegetation types.

Understanding the temporal dynamics of photosynthesis carbon uptake as well as its relationships with the changing climate is challenging. Previous efforts used observations (either in situ [5] or remotely sensed [6]) to establish predictive relationship between photosynthesis rate and climate drivers, and also improved process-based model predictability based on the empirical relationships mined from observations [7]. Machine learning algorithms have been widely used to infer the empirical relationship. For example, Beer 2010 [6] used tree-based regression model to first obtain the predictive relationship between photosynthesis rate and climate drivers, then upscaled FLUXNET observed gross carbon uptake to generate global terrestrial ecosystem carbon sequestration rate maps. In this study, we used an advanced information theory based machine learning technique to infer the robust non-linear relationship between photosynthesis rate and the climate, which will generate significant insights into predictive modeling of photosynthesis rate.

2. Methodology

2.1 FLUXNET data

Eddy Covariance (EC) technique has been widely used in observing CO2 and water fluxes between land surface and the atmosphere. FLUXNET is a globally distributed observational network based on EC technique. Nowadays, hundreds of FLUXNET sites are established and are well maintained to generate reliable fluxes data covering the major vegetation types [5]. We used the observed photosynthesis rates as well as relevant climate drivers from four representative and high-quality FLUXNET sites: 1) FLX_AT-Neu grass site; 2) FLX_DE-Tha evergreen forest site; 3) FLX_NL-Loo evergreen forest site; and 4) FLX_US-Var grass site. Figure 1 showed the density distribution of observed GPP (Gross Primary Productivity, hereafter we will use GPP in stead of photosynthesis rate), solar radiation (FSDS), precipitation (Prcp), air temperature (T), and longwave radiation (FLDS). Black, blue, green, and red shaded area denoted the previously mentioned four sites, respectively.

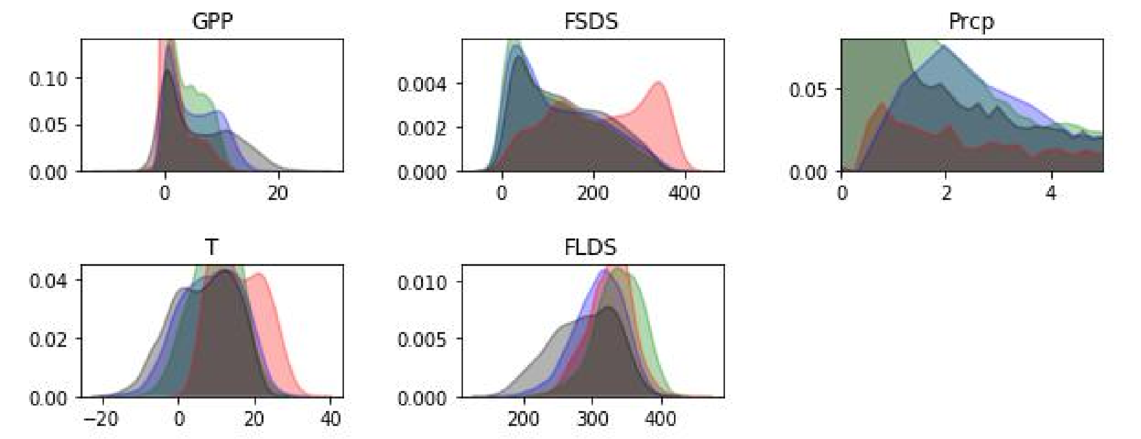

Figure 1. Density distribution of the observed Gross Primary Productivity (GPP), short wave radiation (FSDS), precipitation (Prcp), air temperature (T), and long wave radiation (FLDS). Black, blue, green, and red colors are FLX_AT-Neu, FLX_DE-Tha, FLX_NL-Loo, and FLX_US-Var sites, respecitively

2.2 Mutual information

The responses of Gross Primary Productivity (GPP) to different climate drivers are presumably highly nonlinear, therefore, traditional linear regression based analysis may generally fail. In order to robustly infer non-linear relationship between GPP and the climate drivers, we employed the mutual information concept form information theory. First the Shannon information entropy is defined as [8]:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i) \tag{1}$$

where X is a random variable (in this study could be GPP, temperature and so on), p(x) is the probability density distribution of X. Based on Eqn. 1, we could further define conditional information entropy as:

$$\begin{aligned} &H(X|Y) \\ &= \sum_{y} p(y) H(X|Y=y) \\ &= -\sum_{y}\sum_{x} p(x,y)\log_2 \frac{p(x,y)}{p(y)} \end{aligned} \tag{2}$$

where X, Y are two random variables, p(x), p(y) are marginal distribution of X or Y, p(x,y) is the joint distribution of X and Y. Combine Eqn. 1 and 2, we further derive mutual information as:

$$\begin{aligned} &I(X,Y) \\ &= H(X) - H(X|Y) \\ &= \sum_{y}\sum_{x} p(x,y)\log_2 \frac{p(x,y)}{p(x)p(y)} \end{aligned} \tag{3}$$

where H(X) is information entropy of random variable X, H(X|Y) is conditional information entropy of X to Y (Figure 2). The mutual information represents the level relevance between X and Y.

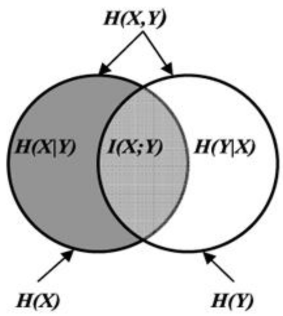

Figure 2. Illustration of mutual information I(X,Y) and its relationship with Shannon information entropy and conditional Shannon information entropy

3. Results and discussion

3.1 Linear coupling

We first analysed possible relationship between the target variable (GPP) and relevant climate drivers (Temperature, precipitation, short wave radiation, and long wave radiation) using linear regression model(Figure 3). The correlation directly informed the strength of linear relationship. We found that GPP was significantly correlated with temperature and shortwave radiation at three out of the four selected sites. FLUX_US-Var site is a exceptional grass land site that GPP was not correlated with any climate variables. Surprisingly, GPP was not correlated with precipitation at any site, given that water availability is one of the most important regulators of vegetation dynamics.

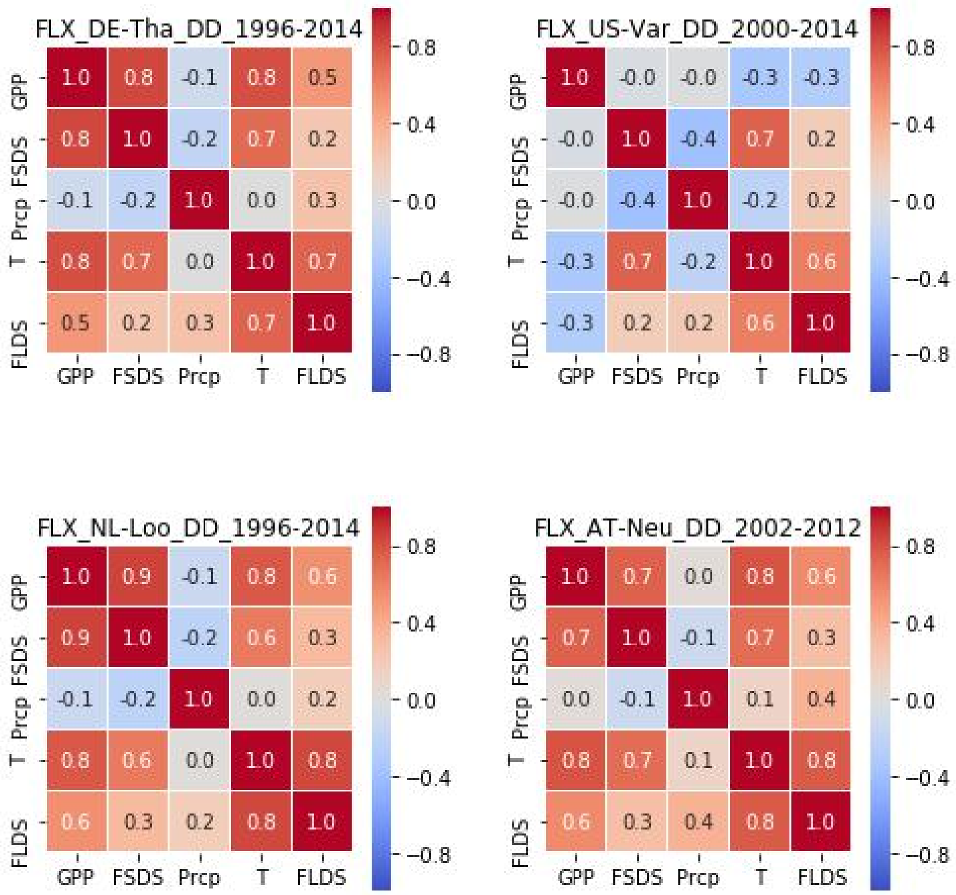

Figure 3. Linear correlation between target variable GPP and climate drivers (T, Prcp, FSDS, FLDS represent air temperature, precipitation, short wave radiation and long wave radiation)

3.2 Non-linear relationship

Given that the vegetation response to climate conditions are highly non-linear, we further analysed the relationships between GPP and climate drivers using the non-linear mutual information metrics, which were depicted in Figure 4. Different color bars represent two probability estimation techniques (see more discussion in section 3.3). In general, high mutual information indicated a strong relationship. We found that precipitation was the major controller of US-Var site GPP, which was not revealed by traditional linear regression model (section 3.1). Overall, short wave radiation and air temperature tended to be important for forest GPP (DE-Tha and NL-Loo), while precipitation tended to be more important at grass site (US-Var and AT-Neu), which is consistent with our theatrical understanding that deep rooting trees are less sensitive to the variation of water availability.

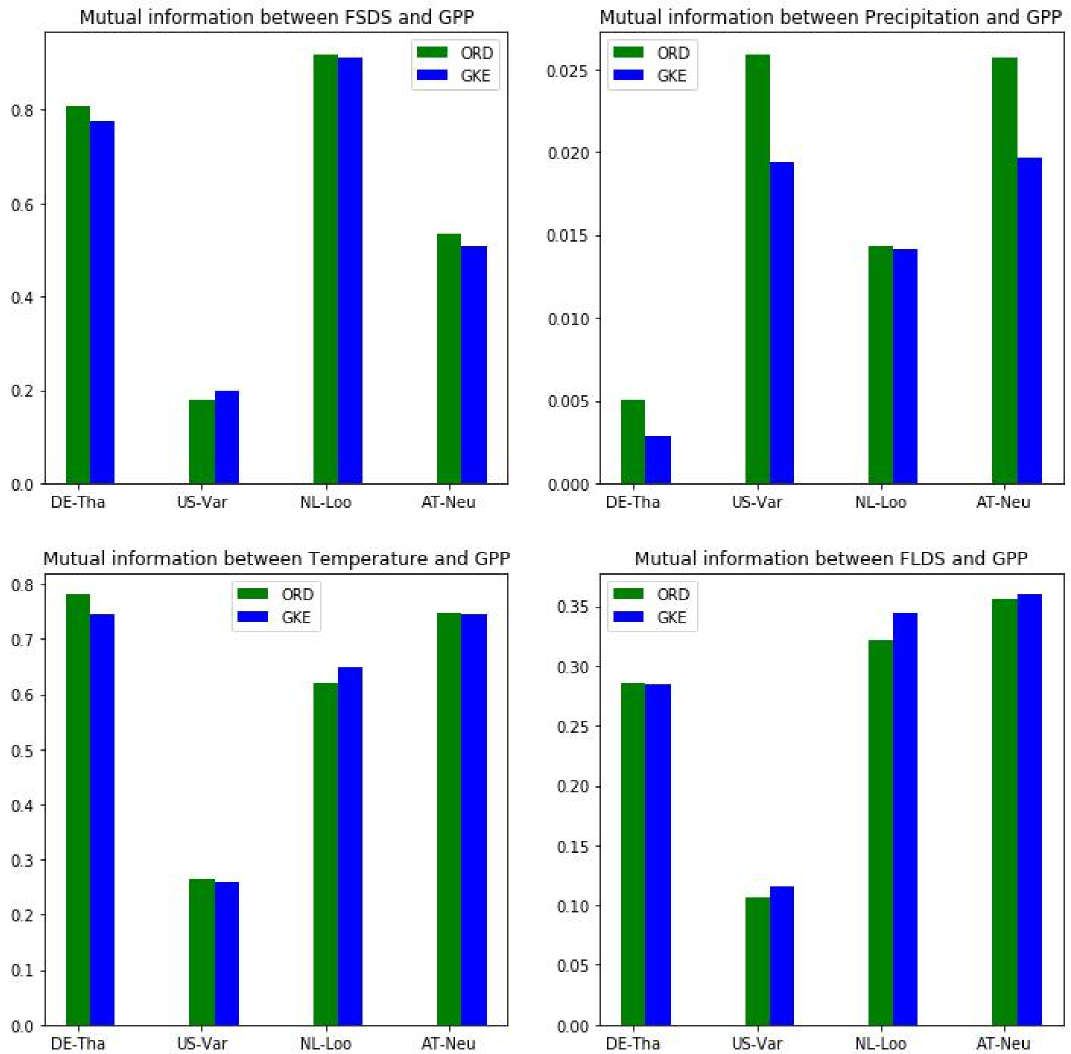

Figure 4. Non-linear relationship between GPP and climate conditions at four sites. Green and blue bars represent different realization of probability estimation.

3.3 Uncertainty analysis

One of the critical uncertainties of mutual information technique is the estimate of Probability Densities Function (PDF). Eqn 1-3 require robust estimation for marginal PDF of X, Y and the joint PDF of X and Y. We first tested classic method of PDF estimation called ordinary ranking (ORD in Figure 4). It divides the variable into equal-distant bins and estimates the density distribution within each bin. Alternatively, we also tested the Gaussian Kernel Estimate (GKE in Figure 4) method that estimate probability density for each bin based on not only the density for the target bin but also considers impacts from adjacent bins given a certain width of influence. Our explorative simulations divided 1D random variable into 8, 9, 10, and 11 bins (Figure 5) and 2D co-varied random variables into 8, 9, 10, and 11 bins (Figure 6 and 7). They consistently showed that GKE method gave a much reliable and consistent estimate of both marginal PDF and joint PDF (Figure 5 red line and Figure 7), while ORD was highly sensitive to the number of bins (Figure 5 blue line and Figure 6). Therefore, the estimate of mutual information between GPP and climate drivers (Figure 4) using GKE (blue bar in Figure 4) should be more reliable.

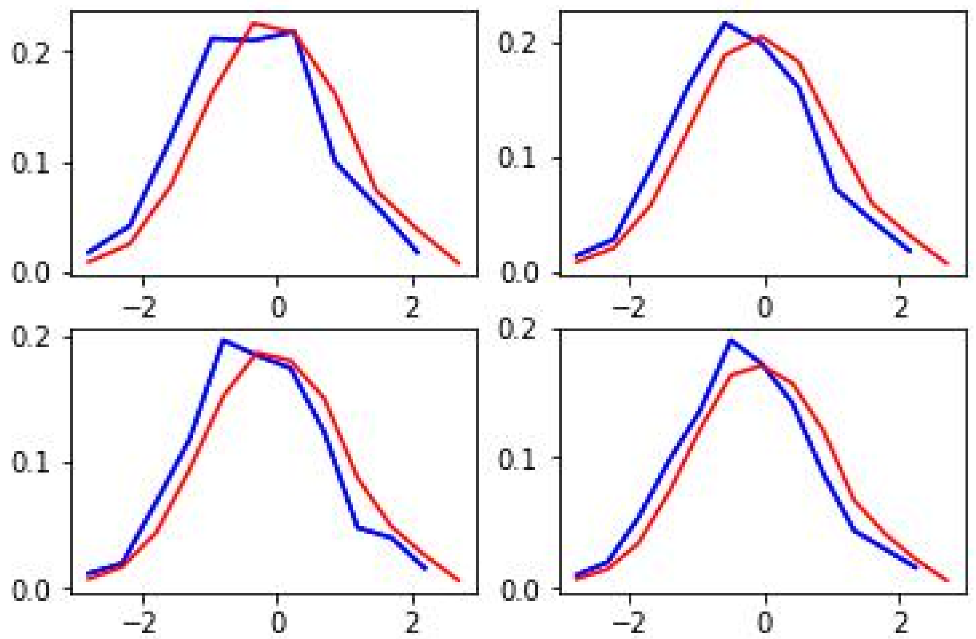

Figure 5. Explorative modelling of PDF with 8 (upper left), 9 (upper right), 10 (bottom left), and 11 (bottom right) bins using ORD method (blue line) and GKE method (red line).

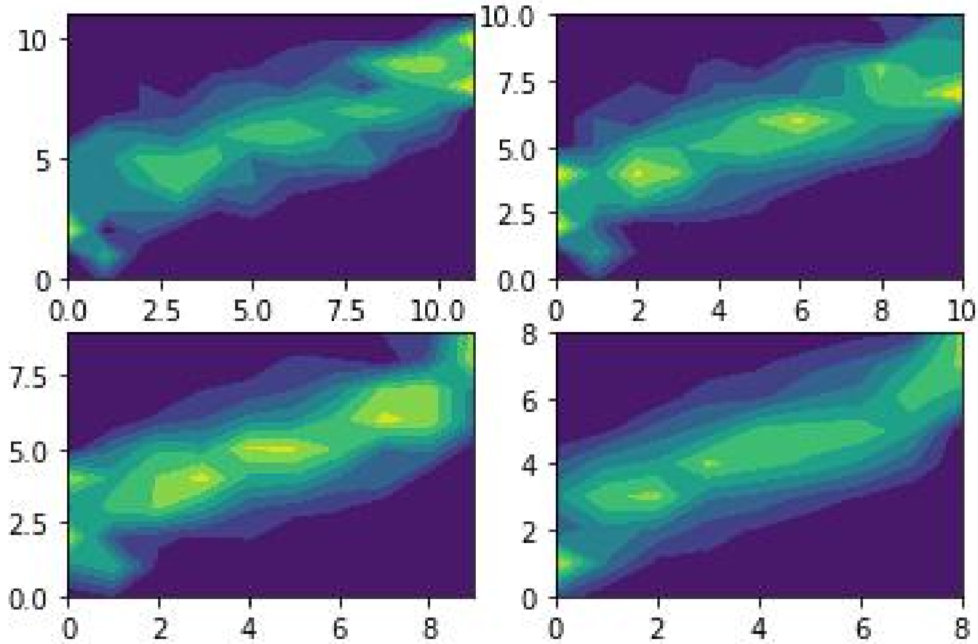

Figure 6. Explorative modelling of joint PDF with 8 (upper left), 9 (upper right), 10 (bottom left), and 11 (bottom right) bins using ORD method

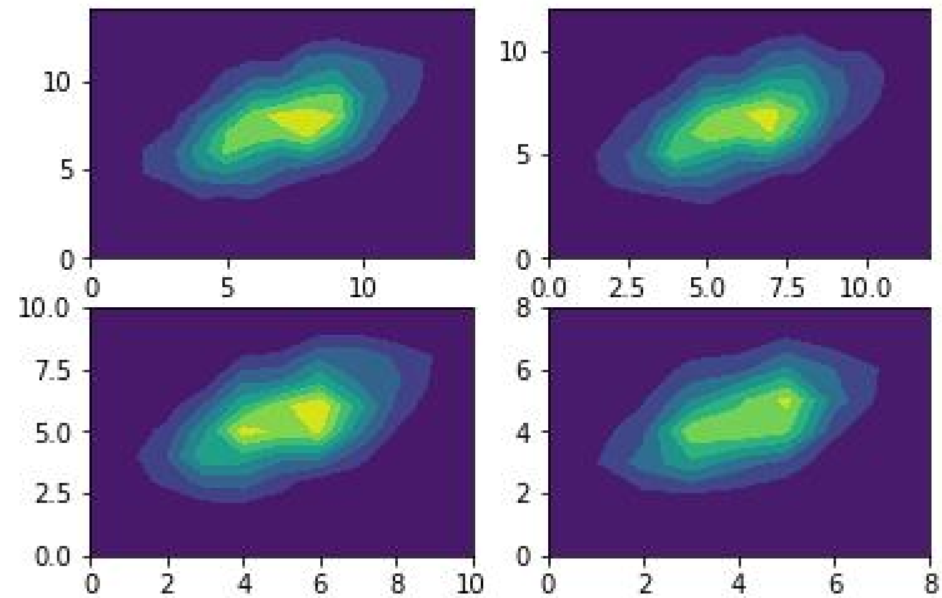

Figure 7. Explorative modelling of joint PDF with 8 (upper left), 9 (upper right), 10 (bottom left), and 11 (bottom right) bins using GKE method

4. Conclusion

Human induced land surface CO_2 emissions and consequently the warming of climate could be partly mitigated by terrestrial plant carbon sequestration via photosynthesis. In this study, we aim to establish possible relationship between plant photosynthesis and climate conditions, thus could help predictive modelling of photosynthesis in the figure. Using mutual information as a proxy of magnitude of relevance between paired variables, we found that temperature and short wave radiation affect photosynthesis significantly at forest sites, while precipitation tended to affected grass land site. We also conducted uncertainty analysis tests using two different density estimate approaches and concluded that GKE is more reliable and consistent than ORD in estimate both marginal and joint PDFs.

5.References

[1] Le Quéré, Corinne, *et al*. 2009 Trends in the sources and sinks of carbon dioxide. *Nature geoscience* **2** 831-36
[2] Feldman, Daniel R, *et al*. 2015 Observational determination of surface radiative forcing by CO2 from 2000 to 2010. *Nature* **519** 339-55
[3] Hansen, J, *et al*., 2010 Global surface temperature change. *Reviews of Geophysics* **48**
[4] Tkemaladze, *et al*., 2016 Climate changes and photosynthesis. *Annals of agrarian science* **14** 119-26
[5] Williams, M, *et al*. 2009 Improving land surface models with FLUXNET data. *Biogeosciences* **6** 1341-59
[6] Beer C, *et al*. 2010 Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science* **329** 834-38
[7] Zhao M, and Running S. W 2010 Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* **329** 940-43
[8] Shannon, C E1948 A Mathematical Theory of Communication. *The Bell System Technical Journal* **27** 379–423, 623–56